REPORT DOCUMENTATION PAGE

AFRL-SR-BL-TR-00-

0666

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, sea... reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including sugg... rmation Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paper...

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPOR...
	November 2000	Final Tech Report 1 Sep 97 to 31 Aug 00

4. TITLE AND SUBTITLE
Averaging Methods Applied to Nonlinear Prediction Equations

5. FUNDING NUMBERS
F49620-97-1-0440

6. AUTHOR(S)
Richard L. Pfeffer, Distinguished Research Professor and
Rossby Professor of Meteorology

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Geophysical Fluid Dynamics Institute
Florida State University
Tallahassee, FL 32306-3017

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Air Force Office of Scientific Research
AFOSR/NM
801 N. Randolph Street, Rm 732
Arlington, VA 22203-1977

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
F49620-97-1-0440

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION AVAILABILITY STATEMENT
Approved for public release; distribution unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

The primary purpose of AASERT Grant F49620-97-1-0440 was to support the research efforts first of Ph.D, degree candidate Scott Applequist (who received his degree in 1999), and later of graduate students Gregory Gabrs and Christopher Werner. Two research components were supported as complementary research to that on AFOSR Grant F49620-96-1-0172. These were as follows: A. Application of Nicolaenco-Mahalov Method to Meteorological Equations. B. Statistical Modification of Numerical Forcasts.

20001208 058

14. SUBJECT TERMS

15. NUMBER OF PAGES
7

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

FINAL REPORT

under

Air Force Office of Scientific Research
AASERT Grant F49620–97–1–0440

Averaging Methods Applied to Nonlinear Prediction Equations

by

Richard L. Pfeffer
Distinguished Research Professor and
Rossby Professor of Meteorology

Principal Investigator

Geophysical Fluid Dynamics Institute
Florida State University
Tallahassee FL 32306–3017

November 2000

A. Introduction

The primary purpose of AASERT Grant F49620–97–1–0440 was to support the research efforts first of Ph.D. degree candidate Scott Applequist (who received his degree in 1999), and later of graduate students Gregory Gahrs and Christopher Werner. Two research components were supported as complementary research to that on AFOSR Grant F49620–96–1–0172. These were as follows:

B. Application of Nicolaenco–Mahalov Method to Meteorological Equations

This research was an effort to extend to a broader class of meteorological prediction problems an averaging method that Basil Nicolaenco and Alex Mahalov of Arizona State University previously applied to the shallow water equations under a separate AFOSR contract. Their method applied to that problem yields a nonlinear equation for the evolution of the low frequency rotational wave (including the full effects of the inertia–gravity waves) and linear equations with variable coefficients for the high frequency inertia–gravity waves (including the full effects of nonlinear interactions with the rotational mode). Numerous discussions with Drs. Nicolaenco and Mahalov led us to believe that this method had great potential for application to more realistic meteorological problems (e.g., the Lorenz equations and the more general forecast equations for baroclinic waves in a stably stratified fluid). Accordingly, with the support of AASERT Grant F49620–97–1–0440 and AFOSR Grant F49620–96–1–0172 we set out to

apply the averaging method to a heirarchy of important meteorological problems, starting with the equations governing the chaotic Lorenz attractor.

Our motivation for beginning with the Lorenz equations was that they have certain important properties in common with the equations governing large–scale atmospheric motions (viz. instability and behavior on both long and short time scales). If we could get something meaningful by applying the averaging technique to the Lorenz equations, it would seem reasonable to proceed next with an attempt to derive equations governing the slow (10–40 day) fluctuations that determine intraseasonal atmospheric behavior.

Our original reason for believing that the methodology might work on the Lorenz equations is that the solution of these equations describes fast oscillations which continue for a long time around one equilibrium point before shifting to oscillate around the other equilibrium point. We conceived of the fast oscillations as being the ones over which we could average in order to get equations governing the longer term oscillation between attractor basins.

Unfortunately, our research lead us to the conclusion that the averaging method is not applicable to the Lorenz equations and, for the same reason, cannot be applied successfully to more general weather forecast problems. The simplest way to understand why this is so is to recognize that, while the procedure works on small oscillations from equilibrium in a single basin of attraction, it cannot work in the case of an attractor, such as the Lorenz attractor, which consists of motions around two different equilibrium points in planes that are at an angle to each other in phase space. We can linearize the equations about **either** of these equilibria, **but not both at the same time**. As soon as the nonlinear trajectory leaves one basin of attraction, the linearization and the averaging break down. To be more specific, suppose we linearized the Lorenz equations

$$dX/dt = -\sigma X + \sigma Y \tag{1}$$

$$dY/dt = r X - Y - X Z \tag{2}$$

$$dZ/dt = -b Z + X Y \tag{3}$$

around the marginal equilibrium point at which the flow bifurcates from steady convection to oscillatory convection. This point is defined by Lorenz as

$$X_o = Y_o = \pm [b(r-1)]^{1/2}\ ; \quad Z_o = r - 1. \tag{4}$$

with $r > 1$. To make matters a little more concrete, we used $\sigma = 16$, $b = 4$, $r = 33.4545....$ These numerical values put us right at the bifurcation point. Given a small perturbation X', Y', Z' around one of the equilibrium points [say, for example, the one using the plus sign in equation (4)], the numerical solution describes a periodic orbit (limit cycle) around that point. Given a large perturbation, the solution describes an attractor going back and forth between the two wings of the Lorenz butterfly. Just as in the case of unstable convection, the solution oscillates around one equilibrium point for a time that is long in comparison with the orbital

period, and then swings over to the other attractor basin and oscillates around the other equilibrium point for a long time before swinging back again.

The linearized system of equations is

$$dX'/dt = -\sigma X' + \sigma Y' \tag{5}$$

$$dY'/dt = (r - Z_0) X' - Y' - X_0 Z' \tag{6}$$

$$dZ'/dt = Y_0 X' + X_0 Y' - b Z'. \tag{7}$$

The Lorenz equations can, therefore, be written in the form

$$d\mathbf{X}'/dt = A\, \mathbf{X}' + \mathbf{N}, \tag{8}$$

where the vector $\mathbf{X}' \equiv (X', Y', Z')^{tr}$, A is the matrix of coefficients in equations (5)–(7) and **N** represents the nonlinear terms. As Dr. Nicolaenco suggested, we performed a change of basis to an eigenvector basis, letting $\mathbf{X}' = T\,\mathbf{V}$, where the vector $\mathbf{V} \equiv (U,V,W)^{tr}$ and T is the matrix of the eigenvectors

$$T \equiv \begin{bmatrix} e_{11} & e_{12} & e_{13} \\ e_{21} & e_{22} & e_{23} \\ e_{31} & e_{32} & e_{33} \end{bmatrix}. \tag{9}$$

Here $\mathbf{e}_1 \equiv (e_{11}, e_{21}, e_{31})^{tr}$ is the eigenvector associated with the real eigenvalue and $\mathbf{e}_2 \equiv (e_{12}, e_{22}, e_{32})^{tr}$, and $\mathbf{e}_3 \equiv (e_{13}, e_{23}, e_{33})^{tr}$ are the real and imaginary parts, respectively, of one of the complex eigenvectors (the other eigenvector being the complex conjugate of this one). This leads to the transformed system of equations of the form

$$\frac{d\mathbf{V}}{dt} = \begin{bmatrix} \eta & 0 & 0 \\ 0 & 0 & \omega \\ 0 & -\omega & 0 \end{bmatrix} \begin{bmatrix} U \\ V \\ W \end{bmatrix} + \tilde{\mathbf{N}}, \tag{10}$$

where, in our case, $\eta = -21.0$ and $\omega = 14.0648$. The nonlinear terms have the form

$$\tilde{\mathbf{N}} = \begin{bmatrix} aU^2 + bUV + cV^2 + dUW + eVW + fW^2 \\ gU^2 + hUV + jV^2 + kUW + lVW + mW^2 \\ nU^2 + pUV + qV^2 + rUW + sVW + tW^2 \end{bmatrix}, \tag{11}$$

where $a = -.093500$, $b = +.280504$, $c = +.094720$, $d = -.194114$, $e = +.065622$, $f = -.059063$, $g = -.386366$, $h = +.035278$, $j = +.047078$, $k = +.450170$, $l = + .238669$, $m = + .219699$, $n = .572133$, $p = - .786633$, $q = + .294723$, $r = + .151721$, $s = - .374660$, $t = - .022278$.

If we let

$$L \equiv \begin{bmatrix} \eta & 0 & 0 \\ 0 & 0 & \omega \\ 0 & -\omega & 0 \end{bmatrix} \tag{12}$$

we may rewrite the transformed equations in the form

$$d\mathbf{V}/dt = L\,\mathbf{V} + \tilde{\mathbf{N}}\,. \tag{13}$$

Frequency analysis of the numerical solution of the original Lorenz equations reveals that X. Y and Z each contain both high and low frequencies. When we integrate (13) with a small initial perturbation, only V and W contain high frequencies. But, when we integrate this same system of equations with an initial perturbation of *moderate size*, all three components display high frequencies, as in the solution of the original Lorenz equations. An assumed solution of the form

$$\mathbf{V}(t) = e^{LT}\,\mathbf{v}(\tau)\ , \tag{14}$$

where $\mathbf{V}(t) \equiv (U,V,W)$ and $\mathbf{v}(\tau) \equiv (\alpha, \beta, \gamma)$, postulates, however, that only V and W contain high frequencies. It is easy to verify this by substituting the matrix

$$e^{Lt} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \omega t & \sin \omega t \\ 0 & -\sin \omega t & \cos \omega t \end{bmatrix}. \tag{15}$$

into (14), in which case it becomes clear that the assumed solution requires that U(t) have only low frequency components. Since the results of the numerical integration of (13) with an initial perturbation of moderate amplitude show that U(t) contains high frequencies, the assumed solution cannot be valid for other than small perturbations around one of the equilibrium points. In order to verify this analysis, we proceeded to go through the formalism of assuming a solution of the form (14) and applying the averaging method. When we did this, the equations governing the slow time variables α, β and γ took the form

$$\frac{d\alpha}{dt} = \tilde{a}\,\alpha + \tilde{b}\,\alpha^2 + \tilde{c}\,\beta^2 + \tilde{d}\,\gamma^2 \tag{16}$$

$$\frac{d\beta}{dt} = \tilde{e}\,\alpha\,\beta + \tilde{f}\,\alpha\,\gamma \tag{17}$$

$$\frac{d\gamma}{dt} = \tilde{g}\,\alpha\,\beta + \tilde{h}\,\alpha\,\gamma. \tag{18}$$

Here $\tilde{a} = -21.00000$, $\tilde{b} = -\tilde{e} = -\tilde{h} = -.093500$, $\tilde{c} = \tilde{d} =.017829$, $\tilde{f} = -\tilde{g} = .618402$. The algebra was accomplished using a symbolic algebra computer program and many of the parts were checked by hand.

When we solved these equations numerically subject to an initially very small perturbation, the solution tended toward constant values of α, β and γ, which corresponds to a limit cycle around the fixed point X_0, Y_0, Z_0, as it should. The implication is that the fast oscillation corresponding to the matrix L takes place around a fixed point that corresponds to our original equilibrium solution [the one corresponding to the plus sign in equation (4)].

Our original hope was that when we solve the same equations with a moderate–size perturbation, we would get a solution for α, β, γ that would describe the longer term tendency for the trajectory to shift from one wing of the Lorenz butterfly to the other. Instead, the procedure fails and the solution blows up, as anticipated from the above discussion of equations (13)–(15). Our interpretation is that, as soon as the initial perturbation is large enough to carry the trajectory outside of the basin of attraction of one of the equilibrium points, the averaging procedure is no longer valid. In particular, $\mathbf{v}(\tau)$ in equation (14) no longer describes the slow time behavior because the products of terms involving e^{LT} and e^{-LT} no longer average to constants.

C. Statistical Modification of Numerical Forecasts

It has been demonstrated by the National Weather Service (NWS) (e.g., Dagastro, V.J. and J.P Dallavalle, 1997) that forecasts by numerical prediction models can be improved by applying statistical corrections to the model output. The methodology is called model output statistics (MOS). All work prior to that reported here was done using linear regression. Owing to the highly nonlinear nature of atmospheric behavior, it seemed reasonable to explore the possibility that the use of nonlinear statistical techniques could yield greater improvements in weather forecasting. Accordingly, with the partial support of AASERT Grant F49620–97–1–0440, as well as a grant from the National Science Foundation, we set out to test the skill of a number of different statistical methodologies and compare the results with those of linear regression.

Specifically, we investigated the use of NGM analyses over the four–year period December 1992 through March 1996 (NCAR archive ds069.5). We used stepwise regression to screen variables from a large pool of potential predictors consisting of those generally used by NWS in MOS predictions, plus predictors that we added based on dynamical considerations and on the experience of Hall (1996), who demonstrated better than average success over a several year period at Dallas/Ft Worth. As a baseline test of skill for comparison with all other methods, we used the same linear regression procedures that are followed by the National Weather Service

with predictors selected from the larger pool by stepwise regression. We compared the results of applying a variety of linear, quasi–linear and nonlinear prediction methods to the prediction of the probability of 24–hour accumulated precipitation exceeding .01, .05 and .10 inches during the cold season.

The methods we tested include linear regression, discriminant analysis, neural networks, logistic regression and a classifier system. Logistic regression, also known as generalized linear modeling, can be considered to be a quasi–linear version of generalized additive modeling. It can also be considered to be a degenerate case of a neural network with no hidden layer, since the transform function used is similar to the "squashing" function in a neural net. The classifier system is a method of artificial intelligence that uses a training set of data to determine "if–then" rules relating a predictand to a prescribed set of predictors, when the number of rules to be learned is specified. These rules are conditions, such as "if the magnitude of a given predictor exceeds a certain threshold, increase the prediction of precipitation probability by a given amount." In order to determine the rules, we used a genetic algorithm, which is a technique for searching for the optimal set of parameters.

We used cross validation in which the relationships between predictors and predictand were determined by training each methodology on three of the four years of data and testing the formulas derived in this way on the data for the remaining year (which then represented an independent data set). For each station, we did this four times, using a different year as the independent data set. To measure the degree of success of each of these probabilistic quantitative precipitation forecasts (PQPFs), we used the Brier Skill Score (BSS), which gives the percent improvement of the prediction over climatology.

We found that we could effectively reduce the pool of potential predictors without loss of skill by using layer averaged values of the predictor variables. This led to a more robust set of variables that showed up as the best predictors at many stations and for all four years investigated by cross validation. Published NWS reports reveal that it is customary in developing MOS equations to include in the set of potential predictors variables measured at individual atmospheric levels (say, for example, the specific humidity, the temperature, various advections, etc. at 1000, 950, 850, 700 and 500 mb). When we applied this methodology, we found that the values of certain variables at one elevation were chosen by stepwise regression as the best predictors for one winter, and that the values of the same variables at a different elevation were chosen as the best predictors for another winter. Similar differences occurred from one station to the next during the same winter. We felt that this is not a robust result and would not hold up in future tests with independent data, but that the physical variables chosen as predictors integrated over a substantial atmospheric layer and over a 24–hour period would be more robust predictors. We tested this hypothesis and found that, indeed, predictive skill did not decrease when we used the much smaller set of vertically averaged variables. The same predictors can then be used at many stations and for all winters.

We trained the 5 different methods as discussed above and made probability forecasts of 24–hour precipitation accumulation exceeding .01", .05" and .10" at 154 stations in the eastern half of the United States from Abiline,Texas to Portland, Maine. We found that we could get respectable predictions with no more than two or three predictors at most stations and up to

seven predictors at a few stations. Although adding more predictors increases the skill of predictions within the dependent data set, it decreases skill when applied to an independent data set. In the following table we compare the mean Brier Skill Scores averaged over all 62 stations obtained using the different statistical methods. It is clear from the table that the scores increase as we progress to larger precipitation amounts, with a much more significant increase occurring between .01" and .05" than between .05" and .10". The table reveals, too, that logistic regression exhibits greater skill than linear regression at all three thresholds, and that discriminant analysis and the classifier system exhibit greater skill than linear regression at the higher thresholds.

Table 2. Comparison of Brier Skill Scores for five methods and three thresholds

	Precipitation Threshold		
	.01"	.05"	.10"
neural network	.339	.444	.473
classifier system	.368	.472	.505
linear regression	.382	.456	.487
discriminant analysis	.372	.482	.514
logistic regression	.393	.496	.521

References

Hall, T., 1996: Brainmaker, A new approach to quantitative and probability of precipitation forecasting. Southern Region Topics, SR/HSD 96–2.

Dagastro, V.J. and J.P Dallavalle, 1997: AFOS–ERA verification of guidance and local aviation/public weather forecasts—No. 23 (October 1994–March 1995). TDL Office Note 97–3